Smoothing of Nonuniformly-Spaced Data

by

Zhenyue Liu

Center for Wave Phenomena
Colorado School of Mines
Golden, Colorado 80401
(303)273-3557

Smoothing of Nonuniformly-Spaced Data

Zhenyue Liu

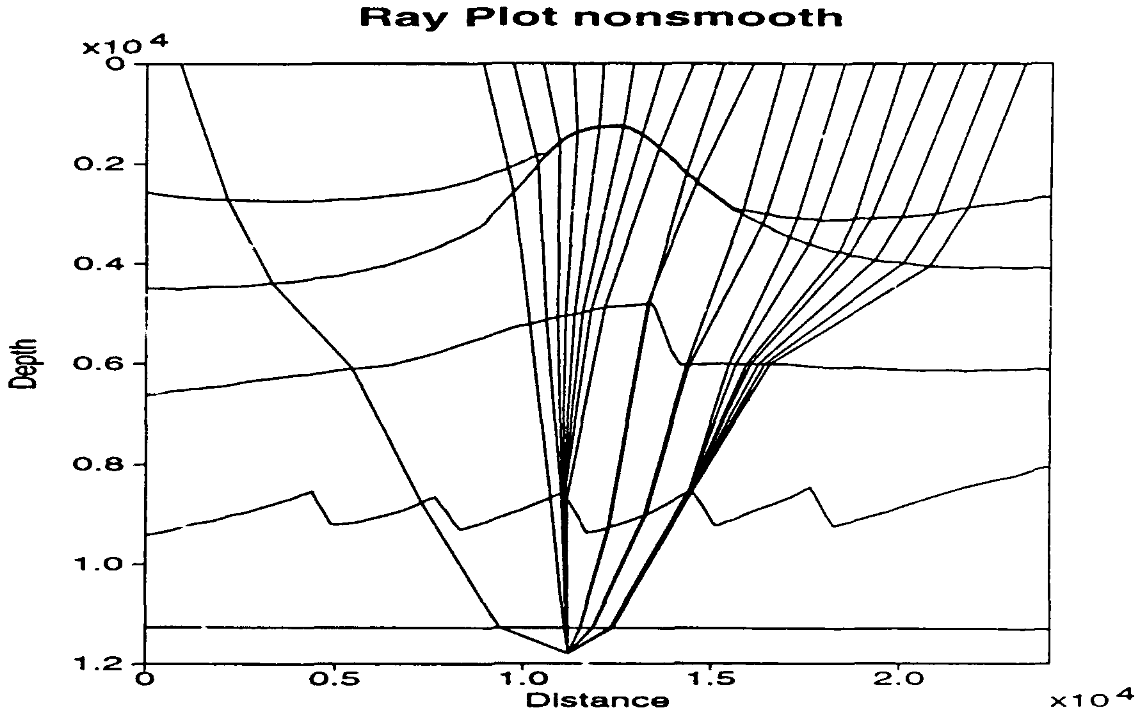
1. INTRODUCTION

Figure 1 shows two outputs from the ray tracing subroutine of the common shot inversion code, CXZCS (Dong, 1989). In both diagrams, the ray from the source to the output point and a suite of rays from the output point to a suite of receivers are shown. The spread of rays in the lower figure (Figure 1b) can be seen to be smoother than the spread in the upper figure (Figure 1a). While the travel times for the two sets of rays is nearly the same, the amplitude-versus-offset distribution for the rays of Figure 1b is also smoother than the amplitude distribution for the rays of Figure 1a.

The model for the Figure 1a was generated by picking points on the interfaces and then connecting them with cubic splines. For Figure 1b, the model of Figure 1a is passed through a smoothing operator which penalizes point distributions that produce large derivatives, while still requires that the modified curve describing each interface still remains "close" to the original model. This paper describes the smoothing operator used to generate the model used in Figure 1b.

The model generating techinique described above is an example of the process of observing and recording data—in that case, (x, z) points of model interfaces. Usually, these data are inaccurate because of noise or measurement error. If we directly connect these data by interpolation to obtain a curve, the result often exhibits a roughness that characterizes the inaccuracies due to the errors, and suggests a fineness of detail that is invalid for the data. When one has a priori information that data represents a smooth phenomenon, then it is important to smooth the observed data consistent with the priori knowlwdge.

Two common smoothing methods are: (1) lowpass Fourier transform and (2) the Moving-average Method. Lowpass Fourier transform is carried out in the frequency domain. This method removes all frequencies above a specific frequency. We can use FFT to speed up the computation, but it requires uniform grid data. The Moving-average Method replaces a value at a point by an average value around that point, but it cannot suppress high frequency components efficiently. In this paper, we give an



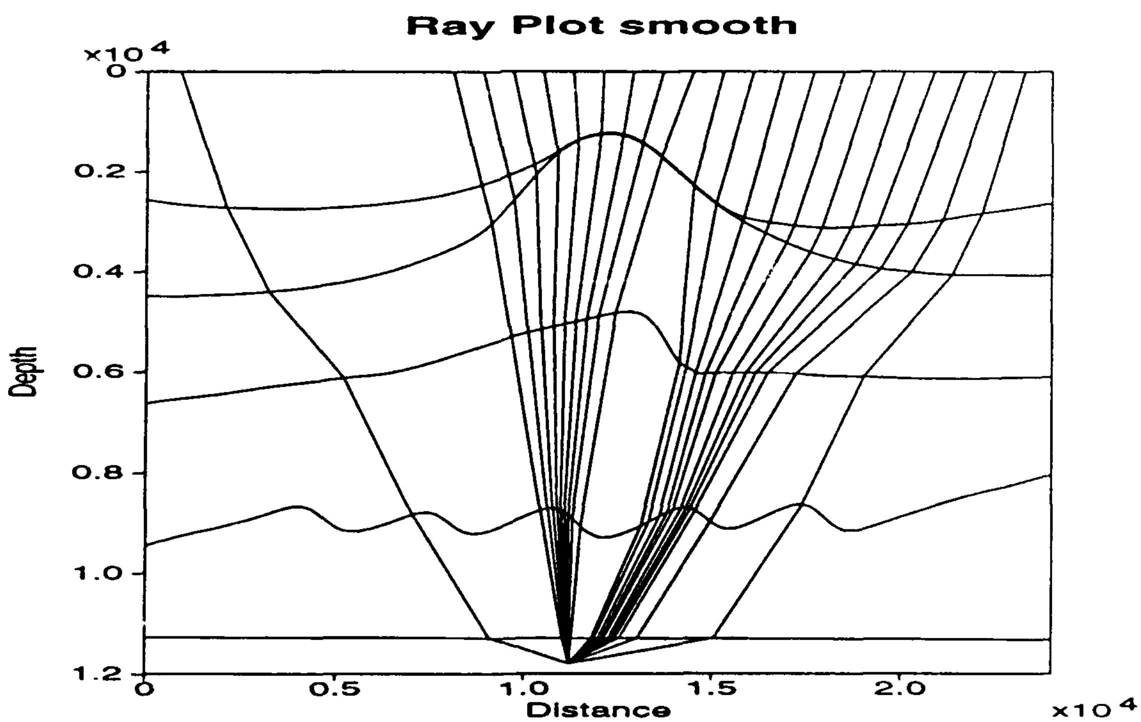


FIG. 1. Ray tracing in the common shot data. The lower one is the smoothed model. The non-smooth model has one ray tracing broken.

alternative smoothing method. A smooth function must have small derivatives. So we hope that a smooth curve is both close to the original one and has as small derivatives as possible. That is, we minimize the weighted squared sum of the error between the desired curve and the original one and some order derivative of the desired curve. This method suppresses high frequency components efficiently instead of removing them. In discretization, this method does not require uniform grid data, and the smooth curve is found by solving a banded linear system whose computational cost is proportional to number of data points.

2. SMOOTHING TECHNIQUE

Mathematical Equation

Let $z = f(x)$ be any continuous function. We solve for a smooth function $g(x)$ that approximates $f(x)$ through the requirement

$$\int [f(x) - g(x)]^2\, dx + \alpha \int (\frac{d^n g}{dx^n})^2\, dx = \min, \tag{1}$$

where $\alpha > 0$ is called the *smoothing parameter* and n is a positive integer. The larger the value of α and the larger n, the smoother $g(x)$ will be.

By using the calculus of variations we can change (1) into a differential equation for $g(x)$:

$$(-1)^n \alpha \frac{d^{2n} g}{dx^{2n}} + g(x) = f(x).$$

This equation has a simple solution by Fourier transform:

$$G(k) = F(k)/(1 + \alpha k^{2n}). \tag{2}$$

Here, we have used capitol letters to denote the Fourier transforms of the corresponding functions denoted by lower case letters in the spatial domain. This expression shows that high-wavenumber components of $f(x)$ are suppressed in the approximation $g(x)$. Equation (2) is same as the Butterworth filter.

Numerical Algorithm

In general, we know a discrete data set, $(x_i, f_i), i = 0, 1, 2, ..., N$, instead of continuous data, $f(x)$. Let $h_i = x_i - x_{i-1}$. We use a finite difference approximation of the differential operator in equation (1). If the h_i's are equal, this work is not difficult. Otherwise, it becomes very complicated for n more than 2. Fortunately, $n = 2$ is a smooth enough index. After discretization equation (1) is equivalent to a linear algebric system of equations for the unknown g_i's. The coefficient matrix of the system is symmetric positive definite, for each n. When $n = 1$ the matrix is tridiagonal; when $n = 2$, the matrix is five-diagonal. Their computations for solving the g_i's are proportional to N.

Choice of smoothing parameter

If the smoothing parameter α is two large, we will lose some low frequency components. Therefore, choosing a suitable smoothing parameter is necessary. We specify a wavenumber k_0, then choose α such that

$$G(k_0) = 0.5F(k_0).$$

Here k_0 should be chosen as some proportion of the Nyquist wavenumber, k_{Nyq}. For example, empirically, I found that taking $k_0 = k_{Nyq}/6$, produced travel times not too far from the traveltimes of the original model, but amplitudes that were much smoother as a function of receiver position than for the unsmoothed model interfaces.

3. COMPARISON OF SMOOTHING METHODS

The Moving-average Method smooths a function as

$$g(x) = \frac{1}{2\Delta x}\int_{x-\Delta x}^{x+\Delta x} f(s)ds, \tag{3}$$

where $f(x)$ is an original function and $g(x)$ is the desired function. In the wavenumber domain, the solution is

$$G(k) = \frac{\sin(k\Delta x)}{k\Delta x}F(k). \tag{4}$$

The larger Δx is, the smoother $g(x)$ will be.

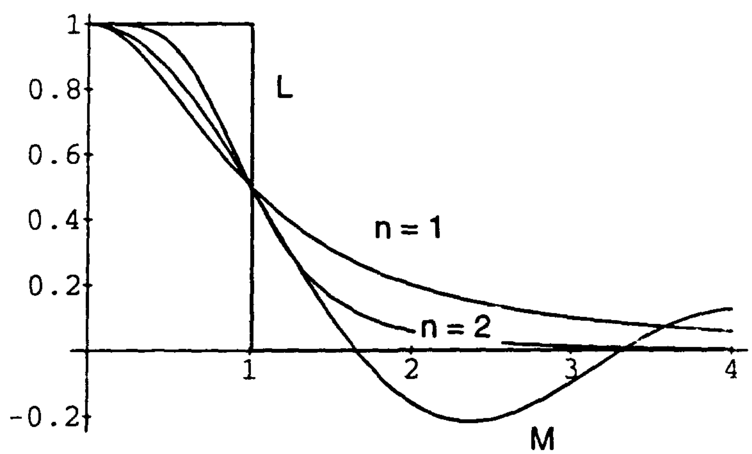


FIG. 2. Smoothing filters in the wavenumber domain. Label M represents the Moving-average method; $n = 1,2$ represent our smoothing method with n as defined in equation (1). L represents the lowpass Fourier transform.

From Figure 2, lowpass Fourier transform can be seen to most effectively eliminate high frequencies, however, at a cost of producing ringing in the spatial domain. With

our method, $n = 2$ provides a sufficiently good approximation of the rectangle of the lowpass Fourier transform while not producing the same undesirable effect in the spatial domain. The (three-term) Moving-average filter decays most slowly in the wavenumber domain. This filter with larger windows and nonuniform sampling is more cunbersome to use than the filter we propose here.

In computational cost, Fourier transform is $O(N \log N)$, the Moving-average is $O(N)$—but with a large constant when decay rate in the Fourier domain is made to compete with our approach, and our method is $O(n^2 N)$.

4. APPLICATION TO RAY TRACING

In seismic data processing, we often meet ray tracing problems. In particular, we consider the medium to consist of constant-velocity layers separated by arbitrary interfaces which are obtained by measurement or digitizing from graphs. To guarantee a successful ray tracing, the modeled interface must be made smooth. Two examples are illustrated in Figure 1 and 3. The smoothed interfaces yield a more uniform coverage by rays, which gurantees a stable amplitude computation (Liu, 1991). Moreover, the smoothed interfaces avoid the ray tracing broken (Figure 1).

5. CONCLUSIONS

In this paper, I introduced a smoothing technique for experimental curve. This method is not restricted by uniform sampling, and smooths a curve efficiently. A key step in practice is how to choose a suitable smoothing parameter. This depends on the properties of the problem to be solved. In the example of my application, I chose the smoothing parameter such that the travaltime is preserved and the amplitude ditribution becomes smoother.

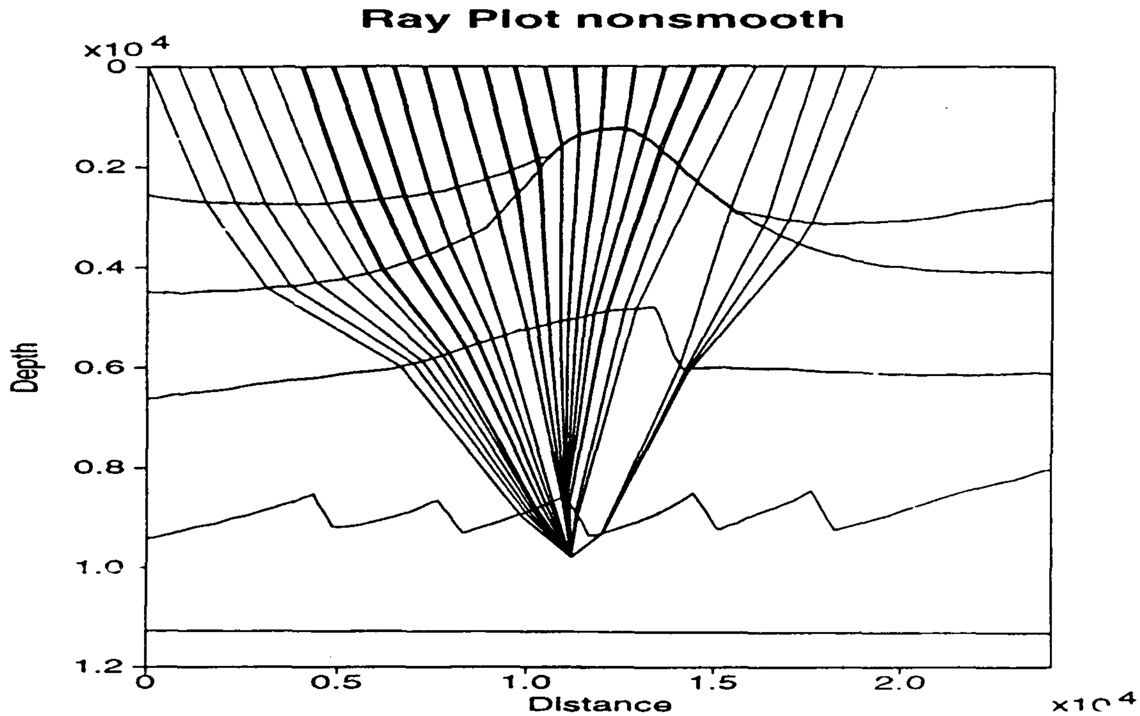
ACKNOWLEDGMENTS

Financial support for this work was provided in part by the Office of Naval Research, Mathematics Division.

REFERENCES

Dong, W., 1989, Finite difference ray tracing and common shot inversion: M.Sc. thesis, Colo. Sch. of Mines.

Liu, Z., 1991, A recursive algorithm for computing ray-parameter derivatives: CWP Project Review.



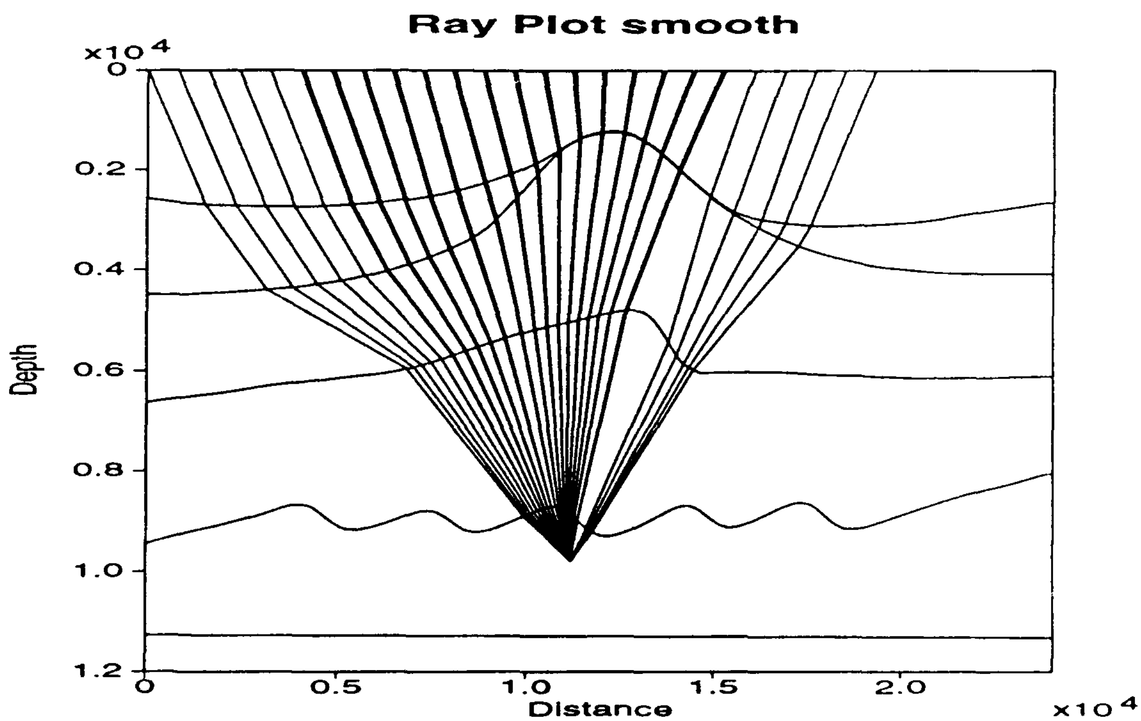


FIG. 3. Ray tracing in the common offset data. The lower one is the smoothed model.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS None.
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Center for Wave Phenomena Colorado School of Mines — 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Mathematical Sciences
6c ADDRESS (City, State, and ZIP Code) Golden, CO 80401	7b ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, VA 22217-5000
8a NAME OF FUNDING / SPONSORING ORGANIZATION — 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-91-J-1267
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO / WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)

Smoothing of Nonuniformly-Spaced Data

12 PERSONAL AUTHOR(S)
Zhenyue Liu

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM 10/1/90 TO 9/30/91	5/20/91	6

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Fourier transform, smoothing parameter, ray tracing Moving-average

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

See Reverse.

Statement A per telecon Dr. Lau
ONR/Code 1111MA
Arlington, VA 22217-5000

NWW 10/2/91

DTIC COPY INSPECTED

Accesion For

NTIS CRA&I ✓
DTIC TAB
Unannounced
Justification

By
Distribution /

Availability Codes

Dist — Avail and / or Special

A-1

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION
22a NAME OF RESPONSIBLE INDIVIDUAL Norman Bleistein	22b TELEPHONE (Include Area Code) (303) 273-3557 — 22c OFFICE SYMBOL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

☆ U.S. Government Printing Office: 1986—607-044

Abstract

This paper introduces a smoothing technique for experimental curves. This method is not restricted by uniform sampling, and smooths a curve efficiently. A key step in practice is how to choose a suitable smoothing parameter. This depends on the properties of the problem to be solved. In the example of application, the smoothing parameter is chosen that preserves the traveltime and the amplitude distribution becomes smoother.